Real-Time High-Resolution Background Matting

By Samuel Golden, Theo Guidroz

Problem Statement

What is Background Matting?

When is it used?

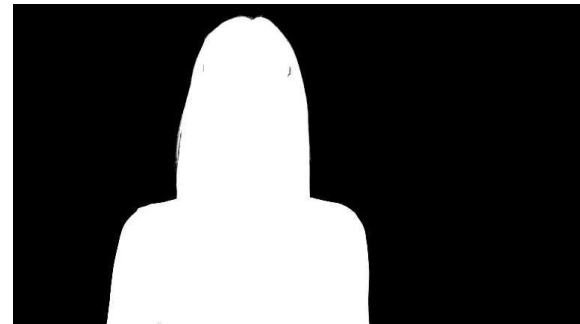Why do we need it?

Input to our models

Background

Source

Alpha

Our Dataset

Dataset publicly available:

- HD background and VideoMatte240K by U of Washington
- Places365 by MIT

What do they consist of?

Can we use it for our project?

Curating the background images

Input to YoloV3
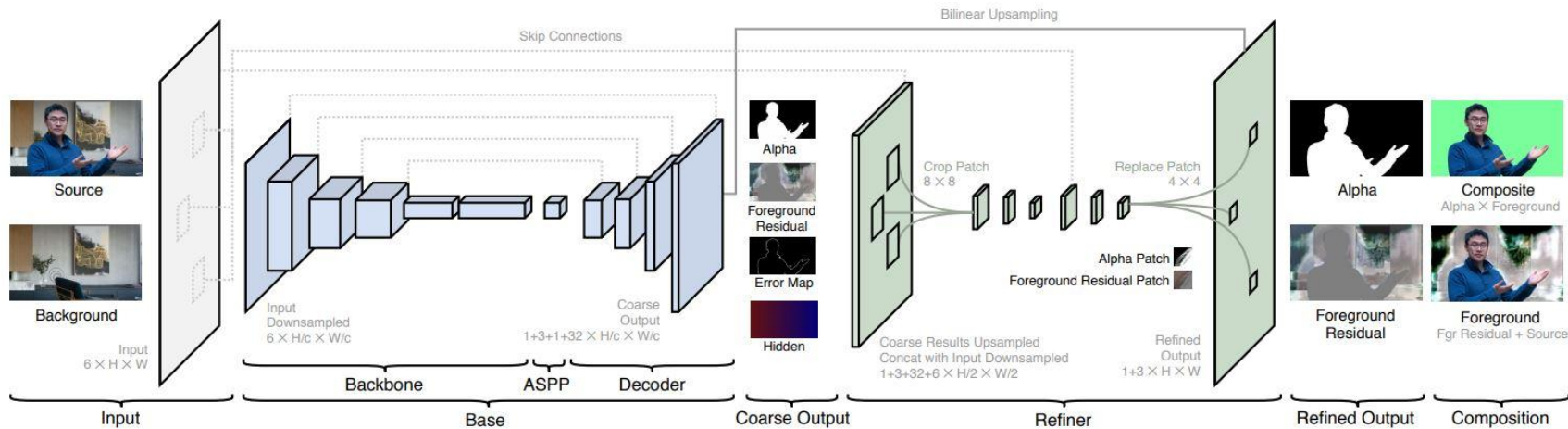
- Person detected
- Image deleted

Input to YoloV3

- No Person detected
- Image kept

# Creating the source images

Model Architecture

Implementation:Base Network

```
resnet = ResNet101(
    weights='imagenet',
    include_top=False,
    input_tensor=img
)
resnet.trainable=False
backbone_in = resnet.input
```

```
# conv block 1
conv1 = Conv2D(ASPP_FILTERS, 1, ...)(x)
conv1 = BatchNormalization(momentum=MOMENTUM, epsilon=EPSILON)(conv1)
conv1 = ReLU()(conv1)

# conv block 2
conv2 = Conv2D(ASPP_FILTERS, 3, ...)(x)
conv2 = BatchNormalization(momentum=MOMENTUM, epsilon=EPSILON)(conv2)
conv2 = ReLU()(conv2)

...

pyr = tf.concat([conv1,conv2,conv3,conv4,pool], axis=-1)
```

Implementation: Refinement Network
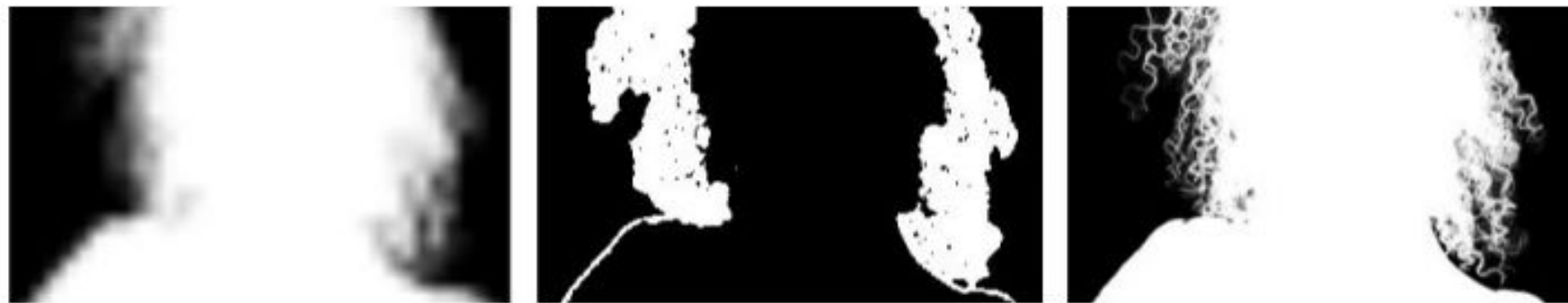
(a) Coarse   (b) Selection   (c) Refined

Implementation: Loss Function

```python
def base_alpha_mse_loss(pha_true, preds):
    pha_pred = tf.clip_by_value(preds[:, :, :, 0:1], 0, 1)
    pha_true = tf.cast(tf.reshape(pha_true, tf.shape(pha_pred)), tf.float32)
    return tf.math.squared_difference(pha_true, pha_pred)


def full_alpha_mse_loss(pha_true, pha_pred):
    pha_true = tf.cast(tf.reshape(pha_true, tf.shape(pha_pred)), tf.float32)
    return tf.math.squared_difference(pha_true, pha_pred)
```

Results

Further improvement

- Train the base network on the entire dataset provided
- Understand why the refinement network crashes during training
- Test is real time

References

[1] A. Farhadi J. Redmon. *YOLOv3: A Incremental Improvement*. 2018. DOI: http://arxiv.org/abs/1804.02767.

[2] Sengupta Lin Ryabtsev. *Real-Time High-Resolution Background Matting*. 2020. DOI: https://arxiv.org/pdf/2012.07810v1.pdf.